

Strict error bounds for linear solid mechanics problems using a subdomain-based flux-free method

Régis Cottureau · Pedro Díez · Antonio Huerta

Abstract We discuss, in this paper, a flux-free method for the computation of strict upper bounds of the energy norm of the error in a Finite Element (FE) computation. The bounds are strict in the sense that they refer to the difference between the displacement computed on the FE mesh and the exact displacement, solution of the continuous equations, rather than to the difference between the displacements computed on two FE meshes, one coarse and one refined. This method is based on the resolution of a series of local problems on patches of elements and does not require the resolution of a previous problem of flux equilibration, as happens with other methods. The paper concentrates more specifically on linear solid mechanics issues, and on the assessment of the energy norm of the error, seen as a necessary tool for the estimation of the error in arbitrary quantities of interest (linear functional outputs). Applications in both 2D and 3D are presented.

Keywords Verification · A posteriori error estimation · Error bounds · Residual-based estimators · Flux-free error estimator

R. Cottureau
International Center for Numerical Methods in Engineering (CIMNE), E. T. S. d'Enginyers de Camins, Canals i Ports de Barcelona, Universitat Politècnica de Catalunya, Barcelona, Spain

Present Address:
R. Cottureau (✉)
Laboratoire MSSMat, CNRS UMR 8579 École Centrale
Paris, Châtenay-Malabry, France
e-mail: regis.cottureau@ecp.fr

P. Díez · A. Huerta
Laboratori de Càlcul Numèric, Departament de Matemàtica Aplicada III, Universitat Politècnica de Catalunya, Barcelona, Spain
e-mail: pedro.diez@upc.edu

1 Introduction

In the past few decades, research and industry in the field of mechanics have relied increasingly on computational tools. The models and the resolution methods have grown increasingly complex and their careful assessment has become unavoidable. In particular, the error arising from the resolution of equations defined on a continuum by the finite element (FE) method has to be estimated and controlled (the so-called “verification” [1]). Hence, this paper describes a technique for the estimation of bounds on the energy norm of that error, in the particular setting of linear solid mechanics.

This technique is an a posteriori error estimation method, which means that it uses the output of the FE computation to assess its accuracy. Three groups of techniques exist within that general class (see [2] for a more detailed review, and [3,4] for recent journal special issues on the subject): one based on the so-called constitutive relation error, by Ladevèze and co-workers (see for example [5]); another based on the comparison of the discontinuous stress field computed by the FE method and a regularized version, following the leading work of Zienkiewicz and Zhu [6]; and, finally, a large family of methods, generically called implicit residual methods, which are based on the (approximate) resolution of a residual error equation on a series of small local problems with appropriate boundary conditions (see for example [7–10], and comparisons between approaches in [11–13]). Among these methods, we distinguish between the hybrid-flux methods (also called equilibrated residual methods), where the local problems are element-based, and the flux-free techniques [7, 13–18], where the subdomains are patches of elements. The advantage of the latter is that the boundary conditions on the local problems are trivial, and that they do not require any flux equilibration. In this paper, we will concentrate on

this type of technique, and we will follow more particularly the approach described in [13, 19].

In most of these subdomain-based methods, although the local error estimation problems are posed on smaller geometrical spaces, the functional spaces involved are still infinite-dimensional. The exact error is therefore usually estimated as the solution of an alternate FE problem, posed on a much larger space than the original FE computation. Hence, the bounds computed are given with respect to a “refined” solution and are valid only asymptotically. However, it is much more interesting, from an engineering point of view, to provide strict bounds, that is to say with respect to the exact error. A dual formulation was proposed to attain that goal in the context of hybrid-flux residual estimators [20, 21], and was later extended to the flux-free error estimators in the case of transient convection-reaction-diffusion problems [22, 23].

This paper describes the extension of these concepts for the derivation of strict bounds of the energy norm of the error in the case of linear solid mechanics problems in the flux-free methodology. Further, specific aspects are discussed and two strategies are investigated for the enhancement of the bounds. Finally, the accuracy of the estimated bounds is shown on several examples in both 2D and 3D problems.

2 Model problem and main notations

2.1 Model problem

We consider an elastic polygonal domain $\Omega \subset \mathbb{R}^d$ (see Fig. 1). The boundary, $\Gamma = \partial\Omega$, is divided into two complementary disjoint parts Γ^N ($\Gamma^N \neq \emptyset$) and Γ^D , where essential and Neumann boundary conditions are imposed, respectively. The boundary value problem to be solved reads: find $\mathbf{u} : \Omega \rightarrow \mathbb{R}^d$, such that

$$\begin{cases} \text{Div}_x \sigma(\mathbf{u}) + \mathbf{f} = \mathbf{0} & \text{in } \Omega \\ \sigma(\mathbf{u}) \cdot \mathbf{n} = \mathbf{g} & \text{on } \Gamma^N, \\ \mathbf{u} = \mathbf{0} & \text{on } \Gamma^D \end{cases} \quad (1)$$

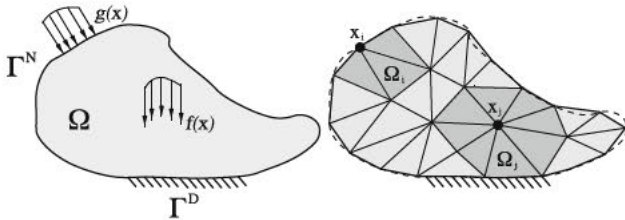


Fig. 1 Model of the problem on the domain Ω , with internal loads \mathbf{f} , boundary forces \mathbf{g} on Γ^N and blocked on Γ^D (left); and corresponding finite element triangulation \mathcal{T}_H with two stars in darker tones, Ω_i and Ω_j , corresponding to nodes \mathbf{x}_i and \mathbf{x}_j , respectively (right)

where the internal force per unit volume $\mathbf{f} \in [\mathcal{H}^{-1}(\Omega)]^d$ and the Neumann boundary tractions $\mathbf{g} \in [\mathcal{H}^{-1/2}(\Gamma^N)]^d$ are given, $\mathcal{H}^\alpha(\mathcal{A})$ is the standard α -Sobolev space over \mathcal{A} , and \mathbf{n} is the outgoing normal vector. $\epsilon(\mathbf{v})$ and $\sigma(\mathbf{v})$ are the strain and stress tensors related to a displacement field \mathbf{v} .

We define $\mathcal{V} = \{\mathbf{v} \in [\mathcal{H}^1(\Omega)]^d, \mathbf{v}|_{\Gamma^D} = \mathbf{0}\}$, the space of admissible fields, and the weak formulation of problem (1) states: find $\mathbf{u} \in \mathcal{V}$, such that

$$a_\Omega(\mathbf{u}, \mathbf{v}) = \ell(\mathbf{v}), \quad \forall \mathbf{v} \in \mathcal{V}, \quad (2)$$

where a_Ω is given by $a_\Omega(\mathbf{w}, \mathbf{v}) = \int_\Omega \sigma(\mathbf{w}) : \epsilon(\mathbf{v}) \, d\Omega$ and ℓ by $\ell(\mathbf{v}) = \int_\Omega \mathbf{f} \cdot \mathbf{v} \, d\Omega + \int_{\Gamma^N} \mathbf{g} \cdot \mathbf{v}|_{\Gamma^N} \, d\Gamma$. The bilinear form $a_\Omega(\cdot, \cdot)$ induces the definition of the energy norm by $\|\mathbf{v}\|_\Omega^2 = a_\Omega(\mathbf{v}, \mathbf{v})$.

The Lax-Milgram theorem ensures that Eq. (2) has a unique solution in \mathcal{V} , that we will denote \mathbf{u}_{ex} . However this solution is usually not available analytically, and we may use the FE method to compute approximations of it.

2.2 Finite Element solution and residual error equation

We therefore introduce a triangulation, \mathcal{T}_H , of Ω , whose elements and vertices are denoted, respectively, $\{T_i\}_{1 \leq i \leq N_e}$ and $\{\mathbf{x}_i\}_{1 \leq i \leq N_v}$, with N_e and N_v the number of elements and vertices. We assume that the triangulation is such that Γ^D and Γ^N consist of entire faces (edges in 2D) of the elements of \mathcal{T}_H . For reasons that will appear in the sequel (more specifically for the solvability of some local problems during the computation of the error bounds, see Sect. 4.2), the FE approximation space over this triangulation is taken to be quadratic, that is $\mathcal{V}_H = \{\mathbf{v} \in \mathcal{V}, \mathbf{v}|_{T_k} \in [\mathbb{P}_2(T_k)]^d, 1 \leq k \leq N_e\}$, where $\mathbb{P}_n(T_k)$, $n \in \mathbb{N}$, is the space of polynomials of order at most n over T_k .

The approximation of \mathbf{u}_{ex} in \mathcal{V}_H is then denoted \mathbf{u}_H , and defined as the unique solution of

$$a_\Omega(\mathbf{u}_H, \mathbf{v}) = \ell(\mathbf{v}), \quad \forall \mathbf{v} \in \mathcal{V}_H. \quad (3)$$

The goal of error estimation techniques is to obtain information on the error $\mathbf{e} = \mathbf{u}_{\text{ex}} - \mathbf{u}_H \in \mathcal{V}$. By linearity, this error is the solution of

$$a_\Omega(\mathbf{e}, \mathbf{v}) = \ell(\mathbf{v}) - a_\Omega(\mathbf{u}_H, \mathbf{v}) =: R(\mathbf{v}), \quad \forall \mathbf{v} \in \mathcal{V}, \quad (4)$$

where the residual R is such that $R(\mathbf{v}) = 0, \forall \mathbf{v} \in \mathcal{V}_H$.

However, the resolution of Eq. (4) is of the same complexity as that of Eq. (2), because \mathbf{e} and \mathbf{v} are still both in the same infinite-dimensional space \mathcal{V} . Therefore, we do not try to solve the problem exactly for \mathbf{e} , but rather look for bounds on the energy norm of \mathbf{e} , namely on $\|\mathbf{u}_{\text{ex}} - \mathbf{u}_H\|_\Omega^2 = \|\mathbf{e}\|_\Omega^2 = a_\Omega(\mathbf{e}, \mathbf{e}) = R(\mathbf{e})$. Such bounds give a global idea of the error resulting from the approximation of \mathbf{u}_{ex} by \mathbf{u}_H . Another approach, more interesting from an engineering point of view, would consist in looking for bounds on

the error in certain quantities of interest, $|\mathcal{J}(\mathbf{u}_{\text{ex}}) - \mathcal{J}(\mathbf{u}_H)|$, but it can be brought back, using a proper error representation [24–29], to that of estimating bounds for the energy norm (for the original problem and a dual one) so that it will not be discussed any further here.

For problems for which the exact solution is known, one can evaluate the ability of a particular method to provide accurate bounds of the energy norm by using a global effectivity index $\eta = \|\hat{\mathbf{e}}\|_{\Omega} / \|\mathbf{e}\|_{\Omega}$, or its local equivalent $\eta_k = \|\hat{\mathbf{e}}\|_{T_k} / \|\mathbf{e}\|_{T_k}$. Both should be close to 1 when the method provides accurate bounds. Further, since the estimates that are discussed in this paper are strict upper bounds of the energy norm of the exact error, we will always have $\eta > 1$, but no equivalent property for the local η_k .

2.3 Domain decomposition

Due to the discontinuity of the stress $\sigma(\mathbf{u}_H)$, the stress $\sigma(\mathbf{e})$ is continuous within each element of the mesh and discontinuous over the element interfaces, denoted Γ^{int} . More specifically, that stress is in the functional space

$$\mathcal{V}_H^{\text{brok}}(\Omega) = \left\{ \sigma = \{\sigma_{ij}\}_{1 \leq i, j \leq d} \mid \sigma_{ij} \in \mathcal{H}^0(T_k), 1 \leq k \leq N_e \right\},$$

and we define the jump of the traction vector as

$$[\![\sigma \cdot \mathbf{n}]\!] = \begin{cases} \sigma_1 \cdot \mathbf{n}_1, & \text{on } \partial\Omega \\ \sigma_1 \cdot \mathbf{n}_1 + \sigma_2 \cdot \mathbf{n}_2, & \text{on } \Gamma^{\text{int}}. \end{cases}$$

where the traction $\mathbf{t}_k = \sigma_k \cdot \mathbf{n}_k$ is the limit of the traction of the interface.

Further in the paper, we will decompose the domain Ω in patches of elements (see Figs. 1 and 2), called stars, $\{\Omega_i\}_{1 \leq i \leq N_v}$, and defined by $\Omega_i = \{\cup T_k \mid \mathbf{x}_i \in T_k\}$. The star Ω_i , centered on \mathbf{x}_i , is therefore the set of elements in contact with the node \mathbf{x}_i . It should be noted that each element is a member of several stars (namely the stars centered on each of its vertices) so that the patches defined here overlap one another. These stars appear in the derivation of our error bounds because they are the support of the linear FE interpolation functions $\{\phi_i(\mathbf{x})\}_{1 \leq i \leq N_v}$, such that $\phi_i(\mathbf{x}_j) = \delta_{ij}$, and $\sum_{i=1}^{N_v} \phi_i(\mathbf{x}) = 1, \forall \mathbf{x} \in \Omega$. Note that N_v denotes the number of vertices of the mesh, rather than the number of nodes, and that these numbers are equal only when using a linear FE method. Note also that the functions ϕ_i do not necessarily coincide with the functions used for the FE interpolation of \mathbf{u}_H , but that they are in \mathcal{V}_H .

The next section describes the dual formulation introduced to ensure the strict upper bound character of the error estimate, and the following presents the actual computation of this error bound by the flux-free methodology.

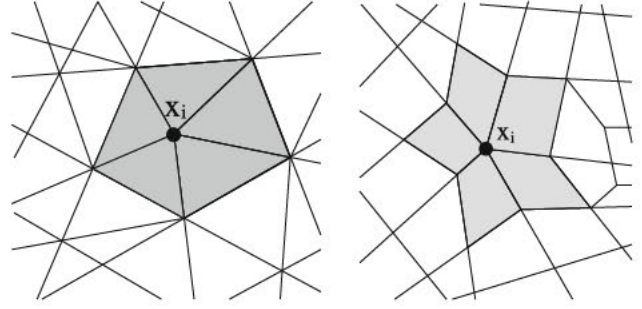


Fig. 2 Examples of stars on a triangular mesh (left) and a quadrilateral one (right)

3 Strict bounds for the energy norm of the error

Following [20,21], strict bounds for the energy norm of the error can be derived by using a dual formulation, so that we start from the strong form of the error estimation problem rather than from the weak form (4).

3.1 Strong form of the residual error equation

Since the stresses corresponding to the FE solution \mathbf{u}_H are possibly discontinuous at the interface between the elements of the mesh, while those corresponding to \mathbf{u}_{ex} are continuous, the stresses corresponding to the error field \mathbf{e} are possibly discontinuous. This has to be taken into account and yields the following strong formulation of the error estimation problem, with jumps in the stresses: find $\mathbf{e} \in \mathcal{V}$, such that

$$\begin{cases} \text{Div}_x \sigma(\mathbf{e}) + (\mathbf{f} + \text{Div}_x \sigma(\mathbf{u}_H)) = \mathbf{0} & \text{in } \Omega \\ \sigma(\mathbf{e}) \cdot \mathbf{n} = (\mathbf{g} - \sigma(\mathbf{u}_H) \cdot \mathbf{n}) & \text{on } \Gamma^N \\ [\![\sigma(\mathbf{e}) \cdot \mathbf{n}]\!] + [\![\sigma(\mathbf{u}_H) \cdot \mathbf{n}]\!] = \mathbf{0} & \text{on } \Gamma^{\text{int}} \\ \mathbf{e} = \mathbf{0} & \text{on } \Gamma^D \end{cases}$$

Extending the definition of \mathbf{g} to $\Gamma^N \cup \Gamma^{\text{int}}$ by setting $\mathbf{g}(\mathbf{x} \in \Gamma^{\text{int}}) = \mathbf{0}$, and remembering the extended definition of the jump operator $[\![\cdot]\!]$ (Sect. 2.3), one gets the equivalent system of equations:

$$\begin{cases} \text{Div}_x \sigma(\mathbf{e}) + \mathbf{f}_H = \mathbf{0} & \text{in } \Omega \\ [\![\sigma(\mathbf{e}) \cdot \mathbf{n}]\!] = \mathbf{g}_H & \text{on } \Gamma^N \cup \Gamma^{\text{int}}, \\ \mathbf{e} = \mathbf{0} & \text{on } \Gamma^D \end{cases},$$

where $\mathbf{f}_H = \mathbf{f} + \text{Div}_x \sigma(\mathbf{u}_H)$ and $\mathbf{g}_H = \mathbf{g} - [\![\sigma(\mathbf{u}_H) \cdot \mathbf{n}]\!]$ are the equivalent loads for the residual error equation.

3.2 Dual formulation of the residual error equation

The dual formulation consists in introducing a new variable \mathbf{q} , representing a stress tensor in $\mathcal{H}_{\text{div}}(\Omega) = \{\mathbf{q} \in L^2(\Omega), \text{Div}_x \mathbf{q} < +\infty \text{ on } T_k, T_k \in \Omega\}$, and verifying

$$\begin{cases} \text{Div}_x \mathbf{q} + \mathbf{f}_H = \mathbf{0} & \text{in } \Omega \\ [\![\mathbf{q} \cdot \mathbf{n}]\!] = \mathbf{g}_H & \text{on } \Gamma^N \cup \Gamma^{\text{int}}. \end{cases} \quad (5)$$

Register for free at <https://www.scipedia.com> to download the version without the watermark

This stress tensor \mathbf{q} is said to be statically admissible, and aims at representing the stress tensor arising from the error, $\sigma(\mathbf{e})$, while relaxing the Dirichlet boundary condition $\mathbf{e} = \mathbf{0}$. We also define, for any stress tensor field \mathbf{q} , the complementary energy as

$$\pi_c(\mathbf{q}) = \frac{1}{2} \int_{\Omega} \mathbf{q} : \mathbb{C}^{-1} : \mathbf{q} \, d\Omega.$$

As stated below, any such statically admissible stress tensor provides an upper bound for the energy norm of the error.

Theorem 1 *Let $\mathbf{e} \in \mathcal{V}$ be a displacement error field, solution of Eq. (4), and $\mathbf{q} \in \mathcal{H}_{\text{div}}(\Omega)$ a stress tensor, solution of the system (5). Then, the latter provides a strict upper bound for the energy norm of the former as:*

$$\|\mathbf{e}\|_{\Omega}^2 \leq 2\pi_c(\mathbf{q}),$$

and the equality is reached for $\mathbf{q} = \sigma(\mathbf{e})$.

Proof For any displacement field $\mathbf{v} \in \mathcal{V}$ cancelling on Γ^D , a stress tensor \mathbf{q} verifying the strong formulation (5) also verifies

$$\int_{\Omega} \mathbf{q} : \varepsilon(\mathbf{v}) = R(\mathbf{v}).$$

As this is the case for the error \mathbf{e} solution of (4), it yields

$$\int_{\Omega} \mathbf{q} : \varepsilon(\mathbf{e}) = R(\mathbf{e}) = a_{\Omega}(\mathbf{e}, \mathbf{e}) = \|\mathbf{e}\|_{\Omega}^2.$$

On the other hand, the expansion of the complementary energy $\pi_c(\mathbf{q} - \sigma(\mathbf{e}))$, which is always positive, and cancels for $\mathbf{q} = \sigma(\mathbf{e})$, yields

$$0 \leq 2\pi_c(\mathbf{q} - \sigma(\mathbf{e})) = 2\pi_c(\mathbf{q}) + \|\mathbf{e}\|_{\Omega}^2 - 2 \int_{\Omega} \mathbf{q} : \varepsilon(\mathbf{e}) \, d\Omega,$$

where we used $\int_{\Omega} \sigma(\mathbf{e}) : \mathbb{C}^{-1} : \mathbf{q} \, d\Omega = \int_{\Omega} \mathbf{q} : \mathbb{C}^{-1} : \sigma(\mathbf{e}) \, d\Omega$, and $\mathbb{C}^{-1} : \sigma(\mathbf{e}) = \varepsilon(\mathbf{e})$. The expected result is obtained by comparison of the last two equations. \square

This theorem provides a way of computing a strict upper bound for $\|\mathbf{e}\|_{\Omega}^2$. However it remains to actually construct a statically admissible stress tensor verifying the system (5). In the form written here, the problem seems only slightly less complex than that of computing directly the exact displacement field \mathbf{u}_{ex} , because it is posed on the entire domain Ω . The next section presents the flux-free method, that allows to replace the global problem (5) by a series of local problems posed on small patches of elements.

4 Computation of strict bounds of the energy norm of the error using the flux-free methodology

Different flux-free methodologies have been developed over the years [7, 13–18], all sharing the property that they somehow make use of a set of functions verifying the partition

of unity property to replace the global problem (either system (5) or a more classical corresponding primal problem) by a set of smaller local problems. They differ by the type of functions and the way in which they are introduced in the equations, and we will here follow the method described in [13, 19].

4.1 The flux-free methodology

In this method, we use the classical linear FE interpolation functions $\{\phi_i\}_{1 \leq i \leq N_v}$ as the set of partition of unity functions and define the local problems as follows: for all vertices i , $1 \leq i \leq N_v$, find a second-order stress tensor $\mathbf{q}^i \in \mathcal{H}_{\text{div}}(\Omega_i)$, such that

$$\begin{cases} \text{Div}_{\mathbf{x}} \mathbf{q}^i + \phi_i \mathbf{f}_H = \mathbf{0} & \text{in } \Omega_i \\ \llbracket \mathbf{q}^i \cdot \mathbf{n} \rrbracket = \phi_i \mathbf{g}_H & \text{on } \Gamma_i \\ \mathbf{q}^i \cdot \mathbf{n} = \mathbf{0} & \text{on } \Gamma^D \cap \partial\Omega_i \end{cases}, \quad (6)$$

where Γ_i is the set of all edges of the star Ω_i , both in its interior and on its boundary, excepted those that might fall on the Dirichlet boundary Γ^D . The homogeneous Neumann boundary condition on the original Dirichlet boundary is added in order to ensure equilibrium of the loads on the star. Note that, using the expanded definitions of \mathbf{f}_H , \mathbf{g}_H and $\llbracket \cdot \rrbracket$, and seeing that ϕ_i cancels on $\partial\Omega_i \setminus (\Gamma^N \cup \Gamma^D)$, the system (6) can also be written

$$\begin{cases} \text{Div}_{\mathbf{x}} \mathbf{q}^i + \phi_i (\mathbf{f} + \text{Div}_{\mathbf{x}} \sigma(\mathbf{u}_H)) = \mathbf{0} & \text{in } \Omega_i \\ \llbracket \mathbf{q}^i \cdot \mathbf{n} \rrbracket = -\phi_i \llbracket \sigma(\mathbf{u}_H) \cdot \mathbf{n} \rrbracket & \text{on } \Gamma_i \setminus \partial\Omega_i \\ \mathbf{q}^i \cdot \mathbf{n} = \phi_i (\mathbf{g} - \sigma(\mathbf{u}_H) \cdot \mathbf{n}) & \text{on } \Gamma^N \cap \partial\Omega_i \\ \mathbf{q}^i \cdot \mathbf{n} = \mathbf{0} & \text{on } \partial\Omega_i \setminus \Gamma^N \end{cases}.$$

Register for free at <https://www.scipedia.com> to download the version without the watermark

This expanded system of local problems on the star is here presented to describe more precisely the different types of boundary conditions that are actually enforced. However, for an easier reading, the rest of the paper will be based on the equations in their reduced form (6).

The flux-free methodology provides an upper bound for the energy norm of the error by the conjunction of Theorem 1 and the following:

Theorem 2 *Let $\{\mathbf{q}^i\}_{1 \leq i \leq N_v}$ be a family of stress tensors, in $\mathcal{H}_{\text{div}}(\Omega_i)$, where each member verifies the system (6) for a star Ω_i . Defining \mathbf{q}_{Ω}^i , $1 \leq i \leq N_v$, as*

$$\mathbf{q}_{\Omega}^i(\mathbf{x}) = \begin{cases} \mathbf{q}^i(\mathbf{x}) & \forall \mathbf{x} \in \Omega_i \\ \mathbf{0} & \forall \mathbf{x} \notin \Omega_i \end{cases},$$

the stress tensor $\hat{\mathbf{q}}$, constructed as

$$\hat{\mathbf{q}} = \sum_{i=1}^{N_v} \mathbf{q}_{\Omega}^i,$$

is then a statically admissible stress tensor, solution of the system (5).

Proof For any point \mathbf{x} in the interior of an element T_k , by definition of the $\mathbf{q}_\Omega^i(\mathbf{x})$, $\hat{\mathbf{q}}(\mathbf{x})$ is obtained as the sum of the $\mathbf{q}^i(\mathbf{x})$, with an index running only on the vertices of that element. Hence $\text{Div}_{\mathbf{x}} \hat{\mathbf{q}} = \sum \text{Div}_{\mathbf{x}} \mathbf{q}^i = -\sum \phi_i \mathbf{f}_H = -\mathbf{f}_H$, because the ϕ_i form a partition of unity. Likewise, for $\mathbf{x} \in \Gamma^{\text{int}}$, $\hat{\mathbf{q}}(\mathbf{x})$ is the sum of the $\mathbf{q}^i(\mathbf{x})$, with an index running only on the vertices of the side (segment in 2D or face in 3D), and $\hat{\mathbf{q}} \cdot \mathbf{n} = \sum \mathbf{q}^i \cdot \mathbf{n} = \sum \phi_i \mathbf{g}_H = \mathbf{g}_H$. \square

Remark 1 The Neumann condition imposed on the Dirichlet boundary of the original problem means that, besides verifying the system (5), the stress tensor $\hat{\mathbf{q}}$ will also verify $\hat{\mathbf{q}} \cdot \mathbf{n}|_{\Gamma^D} = 0$. As $\hat{\mathbf{q}}$ originally aimed at representing the stresses arising from the error $\sigma(\mathbf{e})$, it should rather verify $\mathbf{e}|_{\Gamma^D} = 0$, and possibly $\sigma(\mathbf{e}) \cdot \mathbf{n}|_{\Gamma^D} \neq 0$. Hence, this condition is necessary to impose the equilibrium of the local problems, but is possibly detrimental to the quality of the estimation of the error stress field. However, we will see, in the applications, that it does not seem to deteriorate the quality of the error bounds in a significant manner.

4.2 Solvability of the local problems

Provided that it can be constructed, the family of stress tensors $\{\mathbf{q}^i\}_{1 \leq i \leq N_v}$ described above provides a strict bound of the energy norm of the error. However, the solvability of the local problems (6) remains to be checked. It is given by the following:

Theorem 3 Let \mathbf{u}_H be a FE solution of (3), at least quadratic in the sense that the FE interpolation space \mathcal{V}_H contains at least all element-wise quadratic polynomials over each element T_k .

$$\mathcal{V}_H = \{\mathbf{v} \in \mathcal{V}, \mathbf{v}|_{T_k} \in [\mathbb{P}_2(T_k)]^d, 1 \leq k \leq N_e\}.$$

If the loads \mathbf{f} and \mathbf{g} of the FE problem (3) are piecewise polynomials, of order p_f and p_g , respectively,

$$\mathbf{f} \in \{\mathbf{f}^* \in L^2(\Omega), \mathbf{f}^*|_{T_k} \in [\mathbb{P}_{p_f}(T_k)]^d, 1 \leq k \leq N_e\},$$

and

$$\mathbf{g} \in \{\mathbf{g}^* \in L^2(\partial\Omega), \mathbf{g}^*|_{\partial\Omega_k} \in [\mathbb{P}_{p_g}(\partial\Omega_k)]^d, 1 \leq k \leq N_e\},$$

where $\partial\Omega_k = \partial\Omega \cap \partial T_k$, then, for each star Ω_i , $1 \leq i \leq N_v$, there exists at least one stress tensor \mathbf{q}^i , piecewise polynomial, solution of the system (6).

This theorem is similar to the one proved in [21, Appendix A], but set on stars rather than on elements. Hence, we will base our proof on the latter, and, before proceeding, we recall it here, along with two useful lemmas.

Lemma 1 (Theorem 1 of [21]) For any given set of forcing functions $\mathbf{f}|_{T_k} \in [\mathbb{P}_{p_f}(T_k)]^d$ and $\mathbf{g}|_{T_k} \in [\mathbb{P}_{p_g}(T_k)]^d$, equilibrated in the sense that their equivalent force and moment cancel, that is to say,

$$\int_{T_k} \mathbf{f} d\Omega + \int_{\partial T_k} \mathbf{g} d\Gamma = 0, \quad (7)$$

and

$$\int_{T_k} \mathbf{x} \times \mathbf{f} d\Omega + \int_{\partial T_k} \mathbf{x} \times \mathbf{g} d\Gamma = 0, \quad (8)$$

there exists at least one dual feasible solution $\mathbf{q}_k \in \{\mathbf{q}^* = \{q_{ij}^*\}_{1 \leq i, j \leq d} | \sigma_{ij} \in L^2(T_k), 1 \leq k \leq N_e\}$, verifying

$$\begin{cases} \text{Div}_{\mathbf{x}} \mathbf{q}_k + \mathbf{f} = \mathbf{0}, & \text{in } T_k \\ \mathbf{q}_k \cdot \mathbf{n} = \mathbf{g}, & \text{on } \partial T_k \end{cases}$$

which is piecewise polynomial of degree p , with $p \geq p_g$ and $p > p_f$.

Proof As indicated, this lemma is proved in [21, Appendix A] and will not be derived here. However, it is important to note that the proof requires the splitting of the element T_k . A triangular element is hence split into three, a quadrilateral is split into four triangles, and a tetrahedron into four tetrahedra, each time by adding a node in the center of the original element. Also, the proof has formally been performed only for 2D problems and should be, in a future work, extended to 3D. \square

The two following lemmas will allow us to use the previous one to prove Theorem 3.

Lemma 2 Let \mathbf{u}_H be a FE solution of (3), at least quadratic. Then, for any star Ω_i of Ω , $1 \leq i \leq N_v$, the system of loads $\{\phi_i \mathbf{f}_H, \phi_i \mathbf{g}_H\}$ is self-equilibrated in the sense that the resulting force and moment cancel, that is to say:

$$\int_{\Omega_i} \phi_i \mathbf{f}_H d\Omega + \int_{\Gamma_i} \phi_i \mathbf{g}_H d\Gamma = 0, \quad (9)$$

and

$$\int_{\Omega_i} \mathbf{x} \times \phi_i \mathbf{f}_H d\Omega + \int_{\Gamma_i} \mathbf{x} \times \phi_i \mathbf{g}_H d\Gamma = 0, \quad (10)$$

Proof We first examine the case of the resulting force. As the functions ϕ_i are linear over each element of the mesh, and we considered a quadratic FE interpolation space \mathcal{V}_H , any constant vector \mathbf{a} would fall in \mathcal{V}_H , except that it does not verify the Dirichlet boundary condition on Γ^D . Hence, a modified version of Eq. (3) is verified for $\mathbf{v} = \phi_i \mathbf{a}$, where a term is added to account for the condition on Γ^D . We get, using the fact that the support of ϕ_i is Ω_i , $\phi_i(\mathbf{x} \notin \Omega_i) = 0$,

$$\begin{aligned} & \int_{\Omega_i} \sigma(\mathbf{u}_H) : \varepsilon(\phi_i \mathbf{a}) d\Omega - \int_{\Gamma^D \cap \partial\Omega_i} (\sigma(\mathbf{u}_H) \cdot \mathbf{n}, \phi_i \mathbf{a}) d\Gamma \\ &= \int_{\Omega_i} (\mathbf{f}, \phi_i \mathbf{a}) d\Omega + \int_{\Gamma^N \cap \partial\Omega_i} (\mathbf{g}, \phi_i \mathbf{a}) d\Gamma \end{aligned}$$

Register for free at <https://www.scipedia.com> to download the version without the watermark

Using the divergence theorem, accounting for the possible discontinuity of the stresses $\sigma(\mathbf{u}_H)$ between the elements, and observing that the terms on the Dirichlet boundary condition cancel out, we get that

$$\begin{aligned} & \int_{\Gamma_i} (\llbracket \sigma(\mathbf{u}_H) \cdot \mathbf{n} \rrbracket, \phi_i \mathbf{a}) d\Gamma - \int_{\Omega_i} (\text{Div}_{\mathbf{x}} \sigma(\mathbf{u}_H), \phi_i \mathbf{a}) d\Omega \\ &= \int_{\Omega_i} (\mathbf{f}, \phi_i \mathbf{a}) d\Omega + \int_{\Gamma^N \cap \partial \Omega_i} (\mathbf{g}, \phi_i \mathbf{a}) d\Gamma. \end{aligned}$$

Recalling the definitions of \mathbf{f}_H and \mathbf{g}_H (Sect. 3.1), this equation can be rewritten

$$\left(\int_{\Omega_i} \phi_i \mathbf{f}_H d\Omega, \mathbf{a} \right) + \left(\int_{\Gamma_i} \phi_i \mathbf{g}_H d\Gamma, \mathbf{a} \right) = 0.$$

As this is true for any constant vector \mathbf{a} , we get the expected equation for the resulting force. The same reasoning leads to the equilibration of the moments because, for any constant vector \mathbf{a} , the function $\phi_i \mathbf{a} \times \mathbf{x}$ would be in \mathcal{V}_H , be it for the Dirichlet boundary condition, so that the modified (quadratic) FE equilibrium gives

$$\begin{aligned} & \int_{\Gamma_i} (\llbracket \sigma(\mathbf{u}_H) \cdot \mathbf{n} \rrbracket, \phi_i \mathbf{a} \times \mathbf{x}) d\Gamma - \int_{\Omega_i} (\text{Div}_{\mathbf{x}} \sigma(\mathbf{u}_H), \phi_i \mathbf{a} \times \mathbf{x}) d\Omega \\ &= \int_{\Omega_i} (\mathbf{f}, \phi_i \mathbf{a} \times \mathbf{x}) d\Omega + \int_{\Gamma^N \cap \partial \Omega_i} (\mathbf{g}, \phi_i \mathbf{a} \times \mathbf{x}) d\Gamma, \end{aligned}$$

which leads to the expected result simply by using the rules of the triple product $(\mathbf{a}, \mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}, \mathbf{c}), \forall \mathbf{a}, \mathbf{b}, \mathbf{c}$. \square

Lemma 3 Let $\Omega_i = \bigcup_{k=1}^{N_i} T_k$ be a star, loaded by $\mathbf{f} \in [L^2(\Omega_i)]^d$ and $\mathbf{g} \in [L^2(\Gamma_i)]^d$, equilibrated in the sense of (9) and (10). Then, it is possible to construct a distribution of loads on the boundary of each element $\{\mathbf{g}_k^* \in [L^2(\partial T_k)]^d\}_{1 \leq k \leq N_i}$, such that $\mathbf{f}|_{T_k}$ and \mathbf{g}_k^* be equilibrated, in the sense of (7) and (8), on each element, and

$\llbracket \mathbf{g}^* \rrbracket = \mathbf{g}$, on Γ_i .

Proof The proof follows a rather classical technique, that is described, for example, in [5, Section 8.3] for solid mechanics problems, or in [2, Section 6.4] for scalar problems. We recall here the main steps, in the case of one particular interior star, and send the reader to the above references for a more general proof and details.

We consider an interior star (for which the central node $\mathbf{x}_i \notin \partial \Omega$) with three elements E_1 , E_2 , and E_3 , exterior sides $\Gamma_{1,0}$, $\Gamma_{2,0}$, and $\Gamma_{3,0}$, oriented interfaces $\Gamma_{1,2}$, $\Gamma_{2,3}$, and $\Gamma_{3,1}$ (see Fig. 3, and loads $\mathbf{g}_{1,0}$, $\mathbf{g}_{2,0}$, $\mathbf{g}_{3,0}$, $\mathbf{g}_{1,2}$, $\mathbf{g}_{2,3}$, $\mathbf{g}_{3,1}$ on the corresponding elements. We look for surface loads $\mathbf{g}_{1,2}^*$, $\mathbf{g}_{1,3}^*$, $\mathbf{g}_{2,1}^*$, $\mathbf{g}_{2,3}^*$, $\mathbf{g}_{3,1}^*$, $\mathbf{g}_{3,2}^*$, where the first subscript indicates the element to which the loads refers, and the second the element

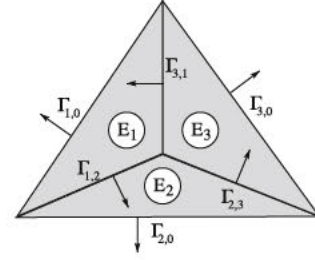


Fig. 3 Example of a star with the notations used in the proof of Lemma 3

that the support interfaces connects to. These loads must verify local jump conditions:

$$\begin{cases} \mathbf{g}_{1,2}^* - \mathbf{g}_{2,1}^* = \mathbf{g}_{1,2} \\ \mathbf{g}_{2,3}^* - \mathbf{g}_{3,2}^* = \mathbf{g}_{2,3} \\ \mathbf{g}_{3,1}^* - \mathbf{g}_{1,3}^* = \mathbf{g}_{3,1}, \end{cases} \quad (11)$$

as well as the equilibrium on each element. For example, on element E_1 , this leads to

$$\begin{aligned} & \int_{\Gamma_{1,2}} (\mathbf{g}_{1,2}^*, \mathbf{a}) d\Gamma + \int_{\Gamma_{1,3}} (\mathbf{g}_{1,3}^*, \mathbf{a}) d\Gamma + \int_{\Gamma_{1,0}} (\mathbf{g}_{1,0}, \mathbf{a}) d\Gamma \\ &+ \int_{E_1} (\mathbf{f}|_{E_1}, \mathbf{a}) d\Omega = 0, \end{aligned}$$

for any solid motion movement \mathbf{a} . Equivalent equations can be written for the other two elements, and, using the jump conditions, this leads to a system of equations in the form

$$\begin{cases} \int_{\Gamma_{1,2}} (\mathbf{g}_{1,2}^*, \mathbf{a}) d\Gamma + \int_{\Gamma_{3,1}} (\mathbf{g}_{3,1}^*, \mathbf{a}) d\Gamma = \ell_1(\mathbf{a}) \\ \int_{\Gamma_{1,2}} (\mathbf{g}_{1,2}^*, \mathbf{a}) d\Gamma + \int_{\Gamma_{2,3}} (\mathbf{g}_{2,3}^*, \mathbf{a}) d\Gamma = \ell_2(\mathbf{a}) \\ \int_{\Gamma_{2,3}} (\mathbf{g}_{2,3}^*, \mathbf{a}) d\Gamma + \int_{\Gamma_{3,1}} (\mathbf{g}_{3,1}^*, \mathbf{a}) d\Gamma = \ell_3(\mathbf{a}) \end{cases}$$

This system allows to compute the value of the moments (the integrals) on the left hand side. Then, a distribution (for example linear) is chosen for the value of $\mathbf{g}_{1,2}^*$, $\mathbf{g}_{2,3}^*$, and $\mathbf{g}_{3,1}^*$, and the last three loads $\mathbf{g}_{2,1}^*$, $\mathbf{g}_{3,2}^*$, and $\mathbf{g}_{1,3}^*$ are derived using Eq. (11). \square

These three lemmas allow to prove Theorem 3 very easily.

Proof of Theorem 3 Indeed, Lemma 2 tells us that the loads on the stars, as defined in system (6), are equilibrated. Hence, Lemma 3 can be used to derive a surface load on the boundary of each element that verifies the conditions required for the application of Lemma 1. The latter gives us the desired result directly.

It should be rightly noted that this theorem does imply that the flux-free method described here cannot be used to derive strict upper bounds for solid mechanics problems that have been solved using linear FE interpolation functions. The rest of this section now concentrates on implementation issues, and, in particular, on the choice of one solution among all the possible ones for each local problem (6).

4.3 Implementation and minimization strategies

We start this section with the derivation of the matrix system corresponding to the system (6). Since it is under-determined, we then turn to the description of two possible ways of choosing one among the possible solutions, and finally discuss several implementation aspects of interest.

Derivation of the matrix system We first choose the order of the polynomial used for the representation of the stress tensor, n , and introduce, for each star, the basis $\mathbf{p}^i = \{\mathbf{p}_m^i\}_{1 \leq m \leq N_0}$ of the (element-wise polynomial) stress tensor space $\{\tilde{\mathbf{q}}^i = [q_{j\ell}]_{1 \leq j, \ell \leq d}, q_{j\ell}|_{T_k} \in \mathbb{P}_n(T_k), T_k \subset \Omega_i\}$. We then introduce the vector $\tilde{\mathbf{Q}}^i = [\tilde{Q}_m^i]_{1 \leq m \leq N_0}$ of coordinates of a stress tensor in that basis as

$$\tilde{\mathbf{q}}^i = \sum_{m=1}^{N_0} \tilde{Q}_m^i \mathbf{p}_m^i.$$

Note that the number N_0 of elements in the basis depends on the number of dimensions d of the physical space, the degree n chosen for the representation of $\tilde{\mathbf{q}}^i$, as well as on the number of elements in the star. For example, for a 2D star of N_e elements, there are $n(n+1)/2$ elements in the polynomial representation and $d(d+1)/2$ independant components of the (symmetric) stress tensor, so that $N_0 = nd(n+1)(d+1)N_e/4$.

We then introduce bases for the divergence space, that is a vectorial basis with a polynomial order $n-1$, and for the tractions on Γ_i and Γ^D . We denote these bases $\{\mathbf{p}_{D,m}^i\}_{1 \leq m \leq N_1}$ and $\{\mathbf{p}_{\Gamma,m}^i\}_{1 \leq m \leq N_2}$, respectively. Considering the same example as above, we would have $N_1 = nd(n-1)N_e/2$, $N_2 = ndN_i$ and $N_3 = ndN_D$, with N_i and N_D the number of sides in Γ_i and Γ^D , respectively. The matrix enforcing the divergence $\mathbf{D} = [D_{j\ell}]$, and those for the tractions $\mathbf{N}^1 = [N_{j\ell}^1]$ and $\mathbf{N}^2 = [N_{j\ell}^2]$ are then defined as

$$D_{j\ell} = - \int_{\Omega_i} (\mathbf{p}_{D-1,j}^i, \text{Div}_{\mathbf{x}} \mathbf{p}_\ell^i) d\Omega, \quad 1 \leq j \leq N_1, \quad 1 \leq \ell \leq N_0,$$

$$N_{j\ell}^1 = - \int_{\Gamma_i} (\mathbf{p}_{\Gamma,j}^i, \llbracket \mathbf{p}_\ell^i \cdot \mathbf{n} \rrbracket) d\Gamma, \quad 1 \leq j \leq N_2, \quad 1 \leq \ell \leq N_0,$$

and

$$N_{j\ell}^2 = - \int_{\Gamma_i} (\mathbf{p}_{D,j}^i, \mathbf{p}_\ell^i \cdot \mathbf{n}) d\Gamma, \quad 1 \leq j \leq N_3, \quad 1 \leq \ell \leq N_0.$$

Finally, we introduce the vectors of coordinates \mathbf{F} and \mathbf{G} of $\phi_i \mathbf{f}_H$ and $\phi_i \mathbf{g}_H$ in the bases \mathbf{p}_{D-1}^i and \mathbf{p}_Γ^i , respectively.

The system (6) can then be written in the following matrix form

$$\begin{bmatrix} \mathbf{D} \\ \mathbf{N}^1 \\ \mathbf{N}^2 \end{bmatrix} [\mathbf{Q}] = \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix}$$

However, this system is, in general, under determined, and a decision must be taken as to which of the possible solutions should be selected. We study here two possibilities, both aiming at minimizing the global complementary energy (associated with $\hat{\mathbf{q}}$) by minimizing the local complementary energy (associated with the \mathbf{q}^i). Indeed, all possible solutions of system (6) provide a strict upper bound for the energy norm of the error through their complementary energy (Theorem 1), so that the ideal choice, for a given order of the polynomial describing the \mathbf{q}^i , would therefore be that which minimizes the complementary energy $\pi_c(\hat{\mathbf{q}} = \sum_{i=1}^{N_v} \mathbf{q}^i)$. This $\hat{\mathbf{q}}$ with the minimum complementary energy is the one that will provide the sharpest bound.

Minimization without accumulation The first natural proposal is therefore to find the \mathbf{q}^i as the solution of the following minimization problem:

$$\mathbf{q}^i = \arg \min_{\tilde{\mathbf{q}}^i} \pi_c(\tilde{\mathbf{q}}^i),$$

under the constraint that the $\tilde{\mathbf{q}}^i$ be solutions of the system (6). This can be enforced through the use of a lagrangian approach. In view of the constraints, we therefore introduce three dual functions λ_1 , λ_2 and λ_3 , and define the lagrangian of the minimization problem as:

$$\begin{aligned} \mathcal{L} = & \frac{1}{2} \int_{\Omega_i} \tilde{\mathbf{q}}^i : \mathbb{C} : \tilde{\mathbf{q}}^i d\Omega - \int_{\Omega_i} (\lambda_1, \text{Div}_{\mathbf{x}} \tilde{\mathbf{q}}^i + \phi_i \mathbf{f}_H) d\Omega \\ & - \int_{\Gamma_i} (\lambda_2, \llbracket \tilde{\mathbf{q}}^i \cdot \mathbf{n} \rrbracket - \phi_i \mathbf{g}_H) d\Gamma - \int_{\Gamma^D \cap \partial \Omega_i} (\lambda_3, \tilde{\mathbf{q}}^i \cdot \mathbf{n}) d\Gamma. \end{aligned}$$

The matrix equations for the system are then derived by cancelling the functional derivatives of that lagrangian \mathcal{L} , with respect to $\tilde{\mathbf{q}}^i$, λ_1 , λ_2 and λ_3 , and projecting the corresponding equations in appropriate bases.

We therefore introduce the mass matrix $\mathbf{M} = [M_{j\ell}]_{1 \leq j, \ell \leq N_0}$, as the projection in the basis \mathbf{p}^i of the complementary energy operator,

$$M_{j\ell} = \frac{1}{2} \int_{\Omega_i} \mathbf{p}_j^i : \mathbb{C} : \mathbf{p}_\ell^i d\Omega, \quad 1 \leq j, \ell \leq N_0,$$

and the vectors \mathbf{L}_1 , \mathbf{L}_2 , and \mathbf{L}_3 of the coordinates of λ_1 , λ_2 , and λ_3 in the bases \mathbf{p}_{D-1}^i , \mathbf{p}_Γ^i , and \mathbf{p}_D^i . The matrix system that

will be solved in each star therefore takes the form:

$$\begin{bmatrix} M & D^T & N^{1T} & N^{2T} \\ D & & & \\ N^1 & & 0 & \\ N^2 & & & \end{bmatrix} \begin{bmatrix} Q \\ L_1 \\ L_2 \\ L_3 \end{bmatrix} = \begin{bmatrix} 0 \\ F \\ G \\ 0 \end{bmatrix} \quad (12)$$

Minimization with accumulation However, although the q^i all minimize locally the complementary energy on their star Ω_i , there is no certainty that $\hat{q} = \sum_{i=1}^{N_v} q^i$ will globally minimize the complementary energy on Ω . Hence, we propose an alternative approach, where all the information that was already computed is used for the minimization. For the first stars that are considered, this process is equivalent to the previous one. However, towards the end of the computation, when most of the stars have been computed already, this method will actually provide the solution that, given all that was already computed, gives the global minimum for the complementary energy of \hat{q} . Note that this means that the order in which the stars are computed will have an impact on the result of the computation. Also, there is no security that the results computed in this manner are better than the ones computed with the previous one. However, we will see in the examples in the next section, that the order variability can actually be used to one's advantage, and that this technique does indeed provide very good results.

We therefore choose an order for the computation and number the stars accordingly $1 \leq i \leq N_v$. When considering the star Ω_i , the local problems on the stars Ω_j , $1 \leq j < i$ have already been solved, and the corresponding q^j computed. We then change the previous minimization problem by the following one:

$$q^i = \arg \min_{\tilde{q}^i} \pi_c \left(\tilde{q}^i + \sum_{j=1}^{i-1} q^j \right), \quad (13)$$

under the same constraints as earlier. The system (12) is then modified in the following way:

$$\begin{bmatrix} M & D^T & N^{1T} & N^{2T} \\ D & & & \\ N^1 & & 0 & \\ N^2 & & & \end{bmatrix} \begin{bmatrix} Q \\ L_1 \\ L_2 \\ L_3 \end{bmatrix} = \begin{bmatrix} -MQ^* \\ F \\ G \\ 0 \end{bmatrix}$$

where the vector Q^* holds the coordinates, in the basis $p(\Omega_i)$, of the stress field obtained from the stars that have already been computed before Ω_i .

Elements refinement In the proof of the existence of a stress tensor verifying the system (6) (Proof of Theorem 3), each element has to be splitted such that two sides of the original element pertain to different elements of the new mesh. Hence triangles are splitted into three new triangles and quadrilaterals into four quadrilaterals, each time by creating a new

node in the interior of the polygon. We chose here to create the new node at the barycenter of the nodes of the element considered. Note that this refinement can be performed at a global level, and once only at the beginning of the computation, since the same element in two different stars has to be splitted each time in the same manner. Note also that this refinement operation is computationally inexpensive but that the cost of the resolution of each local problem (6) rises because the number of elements in each star is larger.

Order of the polynomials The proof of Theorem 3 also indicates that there is a minimal polynomial degree for the stress tensor required to prove its existence, but any higher order polynomial is also valid. Limited numerical experiments were performed and seem to indicate that there is, in most cases, an improvement of the effectivity index when using higher order polynomials. However, this has a very important cost because the number of terms in the polynomial basis increases very rapidly with the polynomial order (in 2D, the second-order basis includes the terms $1, x, y, x^2, xy$ and y^2 , while the third-order basis also includes the terms x^3, x^2y, xy^2 and y^3 ; in 3D, the second-order basis includes the terms $1, x, y, z, x^2, xy, y^2, xz, yz$ and z^2 , while the third-order one is twice as large with the terms $x^3, x^2y, xy^2, y^3, x^2z, xyz, y^2z, xz^2, yz^2, z^3$). The computational cost of the resolution of the systems (6) rises then very rapidly and the improvement of the effectivity index does not seem to balance the cost. In all the applications shown in the next section, the minimum order (quadratic polynomials for the stresses) is always used.

Download the version without the watermark The parallel computation has two advantages: firstly, the computation can be performed very easily on several computers in parallel, and secondly the global computational cost increases only linearly with the number of nodes. Indeed, using the commercial code Cast3m [30] for the FE computations and a Matlab [31] implementation for the error estimation (that could probably be optimized), the tests that we ran indicate that the cost of the error estimation is equivalent to that of the FE computation for a few thousand nodes, and becomes lower for larger computations. Note that when running the error estimation in parallel, either the previous minimization strategy has to be limited to the stars that are actually computed on the local CPU, or some communication between the nodes has to be implemented.

5 Applications

To illustrate the accuracy of the upper bounds developed in this paper, we chose three examples of application: a 2D perforated square plate in plane stress, a 2D gravity dam in plane

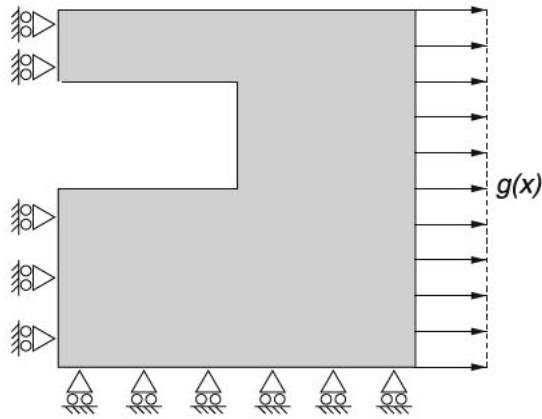


Fig. 4 Model of a quarter part of the perforated square plate

strain, and finally a 3D carabiner. All FE computations are performed using Cast3m [30], and the upper bound computation, using the flux-free method, is implemented in Matlab [31].

5.1 2D perforated square plate in plane stress

The first example that will be considered here is that of a thin square plate with rectangular holes in 2D plane stress. That plate is loaded on the left and right sides by a unit normal traction. The problem is therefore symmetric both in the horizontal and vertical directions so that only a quarter of the plate is actually modeled, with appropriate boundary conditions (see Fig. 4). The Young's modulus and the Poisson's ratio of the material are $E = 1$ and $\nu = 0.3$. Due to the geometry, the solution of this problem is expected to show singularities in the corners of the interior hole, and the numerical errors to be concentrated there. Note that the same problem was already considered in several papers by various authors [13, 19, 21, 32, 33]. We will study here the values of the effectivity indices computed on different meshes, investigate the possibility of using the upper bound of the error to drive an adaptive process, by looking in particular at the maps of local effectivity indices, and comment on the different optimization strategies discussed in Sect. 4.3.

The first mesh that has been considered for this problem is the one on the left in Fig. 5, and is the same as the one used by some of the authors in the references above. The exact value of the displacement field for this problem is not known but can be computed accurately by considering an 'overkill' mesh, that is to say an extremely refined mesh. The displacement field on the crude mesh is then interpolated on the refined one to yield the value of the error and hence its energy norm, either globally or in each element of the crude mesh. As an example, the distribution of the energy norm of the exact error $\|e\|_{\Omega}$ for two different meshes is presented in

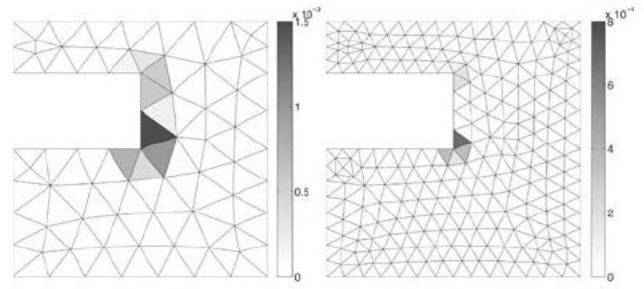


Fig. 5 Two embedded FE meshes with maps of the exact energy norm of the error

Fig. 5. As expected it shows a very strong localization of the error around the corners of the hole.

Following the ideas presented in this paper, and first performing independent minimization on each star, the computation of the upper bound of the energy norm of the error yields a global effectivity index of $\eta = 1.066$ for this first mesh. As it should be as close as possible to 1, and will always remain larger because of the upper bound property, this result is excellent. When using, on the other hand, the accumulation strategy presented in Eq. (13), the order in which the stars are solved becomes important, so that different effectivity indices can be found for the same mesh. We investigate this effect by computing 10,000 times the effectivity index for the same mesh, but changing randomly the order in which the stars are solved. The results are plotted (in light grey shade) as a histogram in Fig. 6. As can be observed by comparison with the value computed without accumulation (the black dot in the figure), the effectivity index computed with the accumulation strategy is always lower with the accumulation strategy (about 10 cases out of 10,000 yield higher effectivity indices, but the values are almost equal to the case without accumulation and are hidden below the dot in the figure). Further, we perform two more series of experiments: on the one hand, 10,000 trials with a random ordering, but always starting by the star with the highest error (the lowest corner of the hole in the plate, see Fig. 5), and on the other hand, 10,000 trials with a random ordering, but always starting by the two stars with the highest errors (both corners of the hole). The results are also plotted on Fig. 6, and show that substantial improvements of the effectivity index can be obtained by performing accumulation starting by the stars that ultimately participate the most in the computation of the upper bound $2\pi_c(\hat{q} = \sum_{i=1}^{N_b} q^i)$. However, note that this can be only used when we have a knowledge a priori of the location of these stars. This may be the case for problems with strong singularities, as is the case here, and in the next problem, but is not true in general. However, in general cases, the accumulation strategy does seem advisable, because it always appears to lower the global effectivity index.

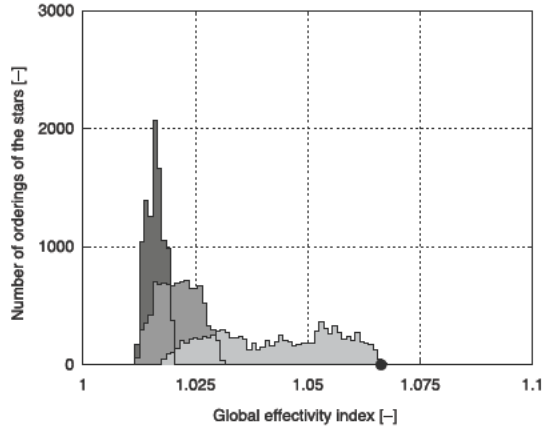


Fig. 6 Influence of the minimization strategy on the global effectivity index: histograms of the effectivity indices obtained using the accumulation strategy with random ordering of the stars (*light grey shade*), starting with the star in the lowest corner of the hole (*grey shade*), and starting with the stars in both corners of the hole (*dark grey shade*); and value obtained without accumulation (*black dot*)



Fig. 7 Convergence of the relative errors $\|e\|_{\Omega}/\|u_H\|_{\Omega}$ and $2\pi_c(\hat{q})/\|u_H\|_{\Omega}$ with the number of DOFs for a uniform (*dashed-dotted line and diamonds*) and an adaptive mesh refinement (*dotted line and crosses*); and upper bounds corresponding to each scheme (*solid lines and circles*)

We then compute upper bounds of the energy norm of the error for a series of embedded meshes to study their rate of convergence. Starting from the mesh on the left of Fig. 5, each triangle is split into four triangles to yield the mesh on the right of the same figure. The process is repeated several times to give a series of embedded meshes with numbers of DOFS ranging from around 500 to 100,000. Note that we are considering TRI6 elements in 2D, so that the number of DOFs is approximately six times that of the number of stars. For example, for the first mesh, there are 74 stars, 255 nodes, and 510 DOFs. For each of these embedded meshes, the upper bound, as well as the exact error (by comparison with an overkill solution), are computed, and both values are

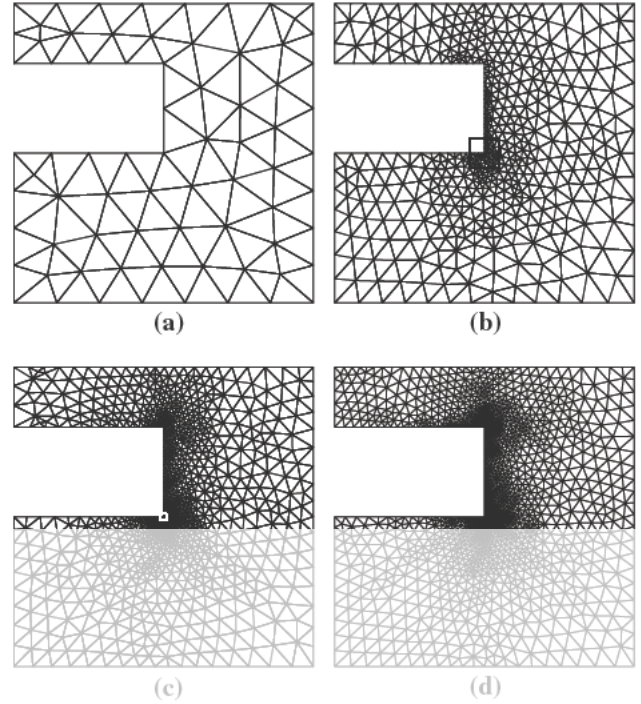


Fig. 8 Sequence of adapted meshes. The errors indicated are the relative energy norms $\|e\|_{\Omega}/\|u_H\|_{\Omega}$ and the black and white squares, respectively in meshes b and c, correspond to the zooms in Fig. 9. a 610 DOFs, error: 6.89%, b 4022 DOFs, error: 1.47%, c 8,754 DOFs, error: 0.78%, d 15,738 DOFs, error: 0.45%

plotted in Fig. 7 (dash-dotted line and corresponding solid line). The global effectivity indices for these meshes range from 1.01 to 1.13.

We finally use the upper bounds to drive an adaptive process, based on a rationale of equidistribution of the energy of the error on the mesh. Hence, the elements with large contribution to the upper bound are split while those with little contribution are merged. More specifically, starting from a given mesh for which the upper bound has been computed, a map of desired element size is drawn by increasing the current element size (by a factor of 1.5) where the local upper bound of the error is small (more specifically, below 20% of the maximum local upper bound), and lowering the current element size (by a factor of 2) where the local upper bound of the error is high (more specifically, over 80% of the maximum local upper bound). That element size map is then fed to a meshing program, EZ4U, to yield a series of adapted meshes, some of which are shown in Fig. 8. The upper bounds and the exact errors (computed using overkill meshes) are plotted in Fig. 7 and show the excellent improvement of the convergence rate obtained with this adaptive process. The effectivity indices for these adapted meshes range from 1.01 to 1.13.

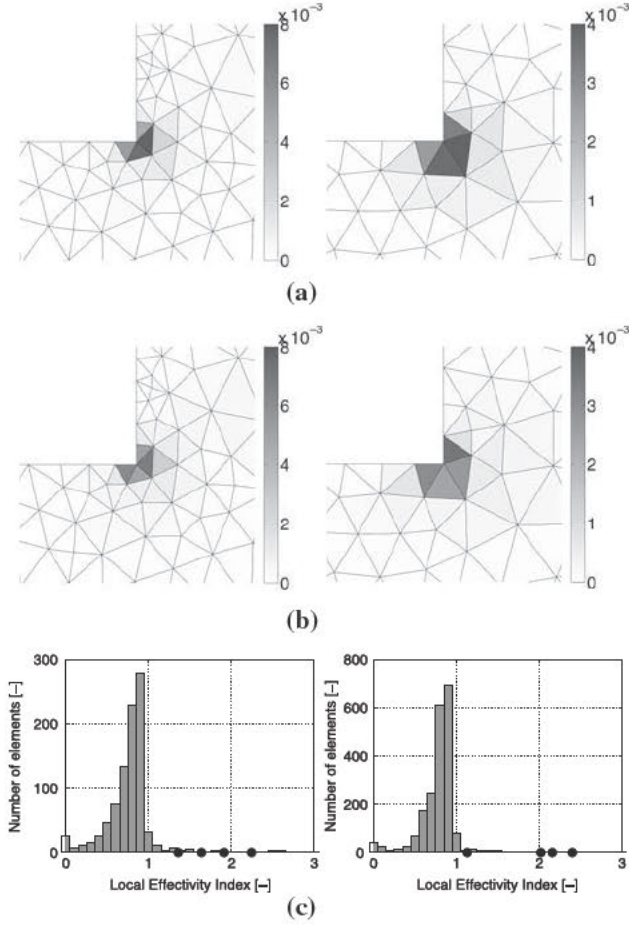


Fig. 9 Distribution maps of **a** the upper bound, computed using the flux-free method, and **b** the exact energy norm of the error, computed using an ‘overkill’ mesh; and **c** histograms of the local effectivity indices. The indices of the four elements with the largest error, summing up to 75% of the total error, are indicated by black dots on the histograms. The upper left figures are zooms in the black box on the mesh in Fig. 8b, and the upper right figures are zooms on the white box on the mesh in Fig. 8c

For two meshes (Fig. 8b and c, respectively), we plot in Fig. 9 (left and right sides, respectively) the local upper bound maps and compare them with the exact error distributions. Although the upper bound property is only ensured at the global level, we observe that the distribution of upper bounds is very close to that of the exact error. This is further shown by plotting the histogram of the local effectivity indices for all elements. It shows a strong clustering, in both cases, a little below 1. However, note that the four main elements, that participate to 75% of the total energy norm of the error (all located around the corners of the hole), all have local effectivity indices above 1 (black dots in the figure). All in all, it seems that the method, at least for this problem with strong singularities, predicts local error slightly below the exact value, while balancing this at the global level by slightly overestimating it for the elements close to the singularity.

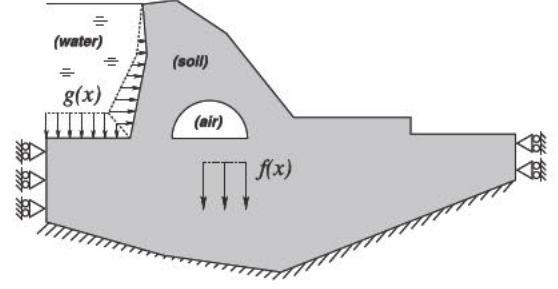


Fig. 10 Model of the gravity dam

5.2 2D gravity dam in plane strain

We then turn to a more elaborate problem, of a 2D gravity dam in plane strain, described in Fig. 10. The model, already used in [6], consists of a dam, with a hole representing a gallery, and loaded both by gravity and a hydrostatic pressure where the water would stand. That last load is approximated by a triangular horizontal load on the (almost) vertical face of the dam (starting at 0 on top of the face, and rising with depth) and a vertical load on the soil standing beneath the reservoir. The boundary conditions below the soil are split into rolling conditions on the left and right sides, and homogeneous Dirichlet condition on the bottom. This model for the boundary conditions therefore allows some degree of settlement of the dam. The Young’s modulus and Poisson ratio are taken as $E = 10$ GPa and $\nu = 0.2$. The total length of the model is $L = 28$ m, for a height of $H = 16$ m from the highest point to the lowest.

The main differences with the previous problem are the complexity of the geometry (several singularity points and a curved boundary), the complexity of the loads (besides the constant surface load as before, a triangular surface load and a volume load), and finally the strong Dirichlet condition on the bottom. We particularly expect the Dirichlet condition on such a large part of the boundary to hinder the quality of our upper bounds since the statically admissible stress tensor field that we estimate will verify a homogeneous Neumann condition on this surface. Further, note that, theoretically speaking, the upper bound property is only valid for piecewise linear boundaries, which is not the case here. However, the overkill meshes will be computed here as refinements of the current one, so that our bounds will indeed be upper bounds with respect to the overkill solutions.

We study here the same issues as in the previous example: quality of the global effectivity index depending on the optimization strategy, rate of convergence of the upper bound in comparison with that of the exact error, and, finally, quality of the local effectivity indices and adaptive process driving.

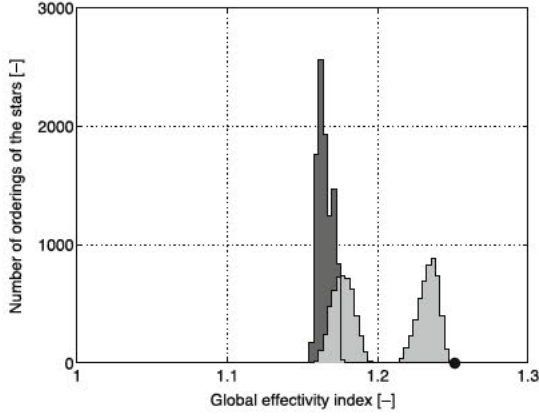


Fig. 11 Influence of the minimization strategy on the global effectivity index: histograms of the effectivity indices obtained using the accumulation strategy with random ordering of the stars (*light grey shade*), and starting with the stars at the geometrical discontinuities (*dark grey shade*); and value obtained without accumulation (*black dot*)

We first plot a histogram of the effectivity indices computed for 10,000 trials (the mesh is that of Fig. 13a), using the accumulation strategy and a random ordering of the stars (see Fig. 11, light grey bars) and compare it with the value obtained without accumulation (black dot on the same figure). As before, the values obtained with accumulation are almost always better than that obtained without accumulation, and when not, they are so close that are hidden on the graph below the black dot. It is interesting to see here that the histogram is bi-modal. By computing, for each star, the correlation (not shown here) between their ranking in the resolution and the value of the effectivity index obtained, it is observed that correlation is extremely high (in negative) for the star in the left corner of the interior gallery and extremely high (in positive) for one just next to it. Further, these values are much higher than for any other star. Hence, it can be concluded that getting an effectivity index in one side or the other of the histogram is mainly driven by the respective ranks of these two stars: taking the star in the corner before means a better effectivity index. We also plot the histogram obtained for 10,000 trials on the same mesh, with accumulation and a random ordering of the stars, but always starting with the star in the left corner of the gallery. As before, the effectivity index is then much better in all trials.

We then study the rate of convergence of the upper bound for a series of homogeneously refined meshes. The first mesh is that of Fig. 13a. Contrarily to the previous example, we do not consider here embedded meshes. Indeed, the geometry of the curved boundary is not well interpolated on that first mesh, and this issue would stick for all meshes derived from it. We therefore used again the meshing program EZ4U, that time asking for a constant element size throughout the mesh. The global effectivity indices obtained for the six meshes considered range from $\eta = 1.15$ to $\eta = 1.35$. These values

are higher than in the previous example, mainly due to the large Dirichlet boundary condition in the bottom. A preliminary correlation study between the rank of the stars on that boundary and the value of the effectivity index seems to indicate the presence of some kind of pattern between the rank of the stars at the singularity points on the Dirichlet boundary condition and their neighbors. However, a more complete study would be required to draw more definite conclusions and possible strategies for improvement of the effectivity index. In any case, the values obtained here are already highly satisfactory.

Finally, we consider the possibility to drive an adaptive refinement process using the upper bound. With the same meshing software as before, and providing local size maps as indicated for the previous example (with an additional smoothing step that is not relevant here), we obtain a series of refined meshes, some of which are plotted in Fig. 13. The global effectivity indices for these meshes also range from $\eta = 1.15$ to $\eta = 1.35$. Note that, in the last mesh, there seems to be irrelevant refinements around the singularities at the bottom of the domain. This is a consequence of the homogeneous Neumann boundary condition satisfied by our estimated stress field at the Dirichlet boundary conditions of the original problem. We therefore get error estimates at these locations that are slightly overestimated, and are therefore refined. Finally, we plot, in Fig. 14, the local maps of the upper bound, next to local maps for the energy norm of the error obtained with the corresponding overkill mesh, as well as the histogram of the local effectivity indices. The conclusion is essentially the same as for the previous example: the local effectivity indices seem to be overestimated for the elements with the highest participation in the global effectivity index, and slightly underestimated for the others. As shown by the rate of convergence obtained using the adapted meshes (Fig. 12), the upper bound seems to be adequate to drive an adaptive process.

5.3 3D carabiner

Finally, we discuss results obtained for a 3D carabiner model. We consider the material to be aluminium with Young's modulus $E = 70$ GPa and Poisson ratio $\nu = 0.3$. The load is a surface load representing the pull of a rope on each side of the carabiner (see Fig. 15) or the loads taking place during an axial stress test. The total force acting on each side is $F = 20$ kN, which is a standard value for the design of such device. The forces acting on each side are equilibrated so that we do not impose any Dirichlet boundary condition and the displacement is defined but for a rigid body motion.

The mesh we consider here has 22908 DOFs (note that the reason for not considering larger meshes is that overkill computations would then not be possible). The global effectivity index for this mesh is very low, $\eta = 1.06$. However,

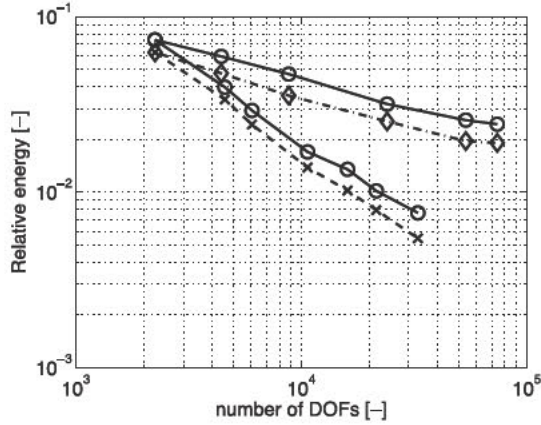


Fig. 12 Convergence of the energy norm of the error with the number of DOFs for a uniform refinement of the mesh (*dashed line and crosses*) and an adaptive refinement (*solid line and circles*)

the authors believe, as was the case in 2D problems, that this effectivity index might increase slightly for problems with strong Dirichlet boundary conditions. The maps of the local upper bounds and energy norm of the exact error (relative values) are then presented on Fig. 16. They show an excellent adequacy between the local estimate and exact value of the energy norm of the error for this problem. Therefore, the flux-free method seems to be, in 3D, an adequate possibility for the driving of adaptive mesh refinement schemes, as was already shown in 2D. The histograms showing the ratio, for each element, of the upper bound on the exact error, are also very good, although they seem slightly more widespread than their counterpart in 2D. Note, however, that there is no singularity in this problem. More study will therefore be needed to evaluate whether the spreading around the mean value for that ratio is due to the dimensionality of the problem or to the singularities in the problem. Further, a large clustering of low values of the ratio can be observed. By comparison of the dark and light grey bars, this clustering is seen to originate from elements for which the error is extremely low (lower than 0.01% of the maximum local value of the exact error), and therefore to be of no importance. As in 2D, we observe here a tendency of the method to slightly underestimate the local effectivity indices for elements with little weight in the evaluation of the global effectivity index and to slightly overestimate those with more importance (see the black dots on Fig. 16).

6 Conclusion

In this paper, we presented a flux-free method for the derivation of strict upper bounds for the energy norm of the error between a FE solution and the exact solution, in linear solid mechanics problems. The computation of the bounds are performed locally on patches of elements, and do not require a

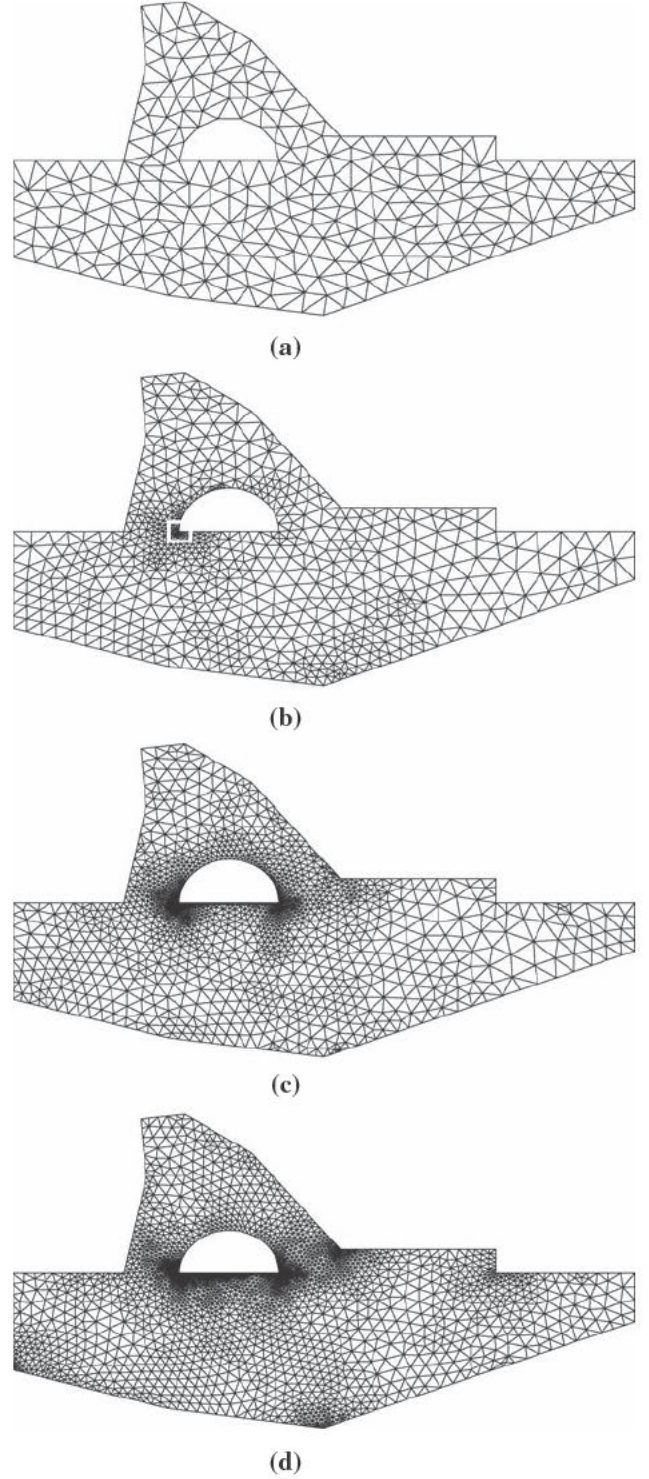


Fig. 13 Sequence of adapted meshes. The errors indicated are the relative energy norms of the exact error and the white box in mesh b corresponds to the zoom in Fig. 14. a 2,248 DOFs, error: 6.25%, b 6,044 DOFs, error: 2.48%, c 16,048 DOFs, error: 1.02%, d 32,892 DOFs, error: 0.55%

previous step of global flux equilibration. The bounds were shown on several examples to be very sharp, and some ways of improving them were discussed, in particular by perform-

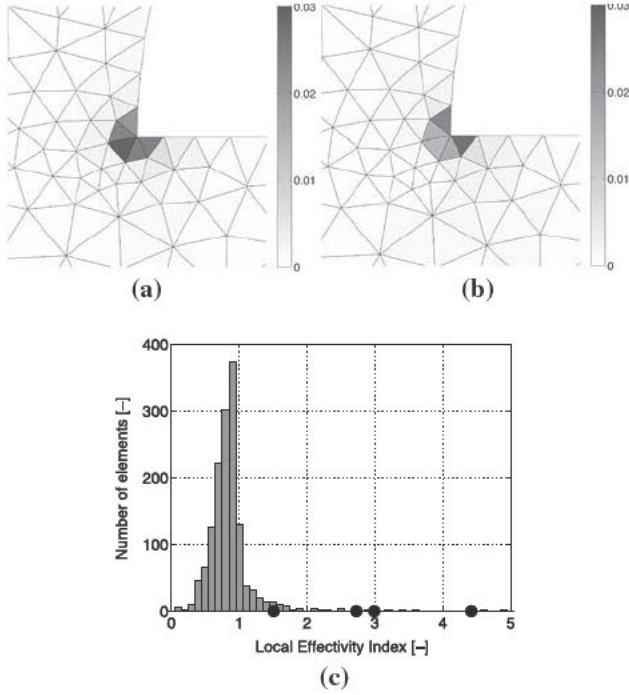


Fig. 14 Distribution maps of **a** the upper bound, computed using the flux-free method, and **b** the exact energy norm of the error, computed using an ‘overkill’ mesh; and **c** histograms of the local effectivity indices. The indices of the four elements with the largest error, summing up to 50% of the total error, are indicated by black dots on the histograms. The mesh figures are zooms in the white box on the mesh in Fig. 13c

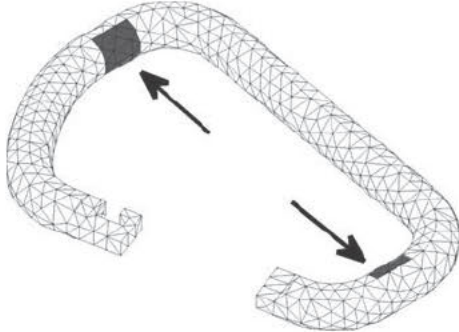


Fig. 15 Model of the carabiner: the arrows indicate the location and direction of the loads

ing accumulation during the minimization, and starting the process by the stars centered on a singularity when it is known. It is also possible to increase the order of the polynomial used to represent the admissible stress field, but this increases the cost of the method disproportionately.

The main limitation of the method we proposed is the requirement to compute the FE solution using quadratic interpolation functions, so as to ensure the equilibration of the loads for the local problems in the error estimation. As discussed, the presence of a large Dirichlet boundary condition seems to decrease slightly the sharpness of the upper bound, and this should be investigated further. It should also be noted

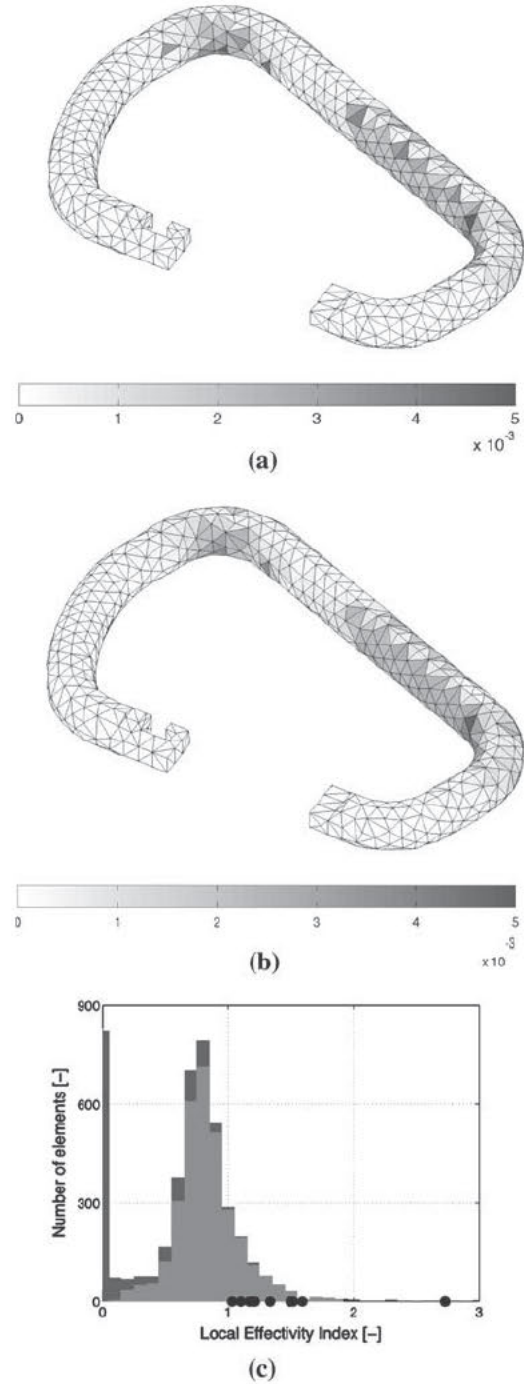


Fig. 16 Distribution maps of **a** the upper bound, computed using the flux-free method, and **b** the exact energy norm of the error, computed using an ‘overkill’ mesh; and **c** histograms of the local effectivity indices, considering all elements (*dark grey bars*) or only those for which the exact error is superior to $1e-4$ times the maximum local exact error (*light grey bars*). The indices of the ten elements with the largest error, summing up to 35% of the total error, are indicated by black dots on the histograms

that, theoretically speaking, the strict upper bound property is true only when considering a boundary which is piecewise-linear.

In the course of this paper, we have shown how to construct a stress field $\hat{\mathbf{q}}$ which is statically admissible in Ω in the sense that it equilibrates exactly all the loads. However, we use here only its complementary energy to yield the desired upper bound. A very promising alternative use of this statically admissible field is to derive strict bounds in the context of nonlinear solid mechanics, using the works of Ladevèze and coworkers [34,35]. This will be the object of a forthcoming paper.

Acknowledgments The authors would like to acknowledge the help of Prof. J. Sarrate and his team at the Laboratori de Càlcul Numèric of the Universitat Politècnica de Catalunya (Barcelona, Spain), and in particular X. Roca, for the help with the use of the program EZ4U, that they develop. Also, we would like to acknowledge fruitful discussions with Prof. H. Ben Dhia (École Centrale Paris) on the appropriate functional spaces for the dual problems.

References

1. Babuška I, Oden JT (2004) Verification and validation in computational engineering and science: basic concepts. *Comp Meths Appl Mech Eng* 193:4057–4066
2. Ainsworth M, Oden JT (2000) A posteriori error estimation in finite element analysis. Pure and applied mathematics. Wiley-Interscience, New York
3. Wiberg NE, Díez P (2006) Adaptive modeling and simulation. *Comp Meths Appl Mech Eng* 195(4–6):205–480
4. Stein E, Stephan EP (2007) Error controlled hp-adaptive FE and FE-BE methods for variational equalities and inequalities including model adaptivity. *Comp Mech* 39(5):555–680
5. Ladevèze P, Pelle JP (2005) Mastering calculations in linear and nonlinear mechanics. Mechanical engineering. Springer, Heidelberg
6. Zienkiewicz OC, Zhu JZ (1987) A simple error estimator and adaptive procedure for practical engineering analysis. *Int J Numer Meth Eng* 24:337–357
7. Babuška I, Rheinboldt WC (1978) Error estimates for adaptive finite element computations. *SIAM J Numer Anal* 15(4):736–755
8. Babuška I, Rheinboldt WC (1978) A posteriori error estimates for the finite element method. *Int J Numer Meth Eng* 12(10):1597–1615
9. Ladevèze P, Leguillon D (1983) Error estimate procedure in the finite element method and applications. *SIAM J Numer Anal* 20(3):485–509
10. Demkowicz L, Oden JT, Strouboulis T (1984) Adaptive finite elements for flow problems with moving boundaries. Part I: variational principles and a posteriori error estimates. *Comp Meths Appl Mech Eng* 46(2):217–251
11. Zhu JZ (1997) A posteriori error estimation, the relationship between different procedures. *Comp Meths Appl Mech Eng* 150(1–4):411–422
12. Choi HW, Paraschivoiu M (2004) Adaptive computations of a posteriori finite element output bounds: a comparison of the ‘hybrid-flux’ approach and the ‘flux-free’ approach. *Comp Meths Appl Mech Eng* 193(36–38):4001–4033
13. Parés N, Díez P, Huerta A (2006) Subdomain-based flux-free a posteriori error estimators. *Comp Meths Appl Mech Eng* 195(4–6):297–323
14. Carstensen C, Funken SA (1999–2000) Fully reliable localized error control in the FEM. *SIAM J Sci Comput* 21(4):1465–1484
15. Machiels L, Maday Y, Patera AT (2000) A ‘flux-free’ nodal Neumann subproblem approach to output bounds for partial differential equations. *Comptes-Rendus de l’Académie des Sci Ser I Math* 330(3):249–254
16. Morin P, Nochetto RH, Siebert KG (2003) Local problems on stars: a posteriori error estimators, convergence, and performance. *Math Comp* 72:1067–1097
17. Prudhomme S, Nobile F, Chamoïn L, Oden JT (2004) Analysis of a subdomain-based error estimator for finite element approximations of elliptic problems. *Numer Meths Partial Diff Eqs* 20(2):165–192
18. Moitinho de Almeida JP, Maunder EAW (2000) Recovery of equilibrium on star patches using a partition of unity technique. *Int J Numer Meth Eng* (accepted). doi:10.1002/nme.2623
19. Parés N (2005) Error assessment for functional outputs of PDE’s: bounds and goal-oriented adaptivity. Ph.D. thesis, Universitat Politècnica de Catalunya, Barcelona, Spain
20. Sauer-Budge AM, Bonet J, Huerta A, Peraire J (2004) Computing bounds for linear functionals of exact weak solutions to Poisson’s equation. *SIAM J Numer Anal* 42(4):1610–1630
21. Parés N, Bonet J, Huerta A, Peraire J (2006) The computation of bounds for linear-functional outputs of weak solutions to the two-dimensional elasticity equations. *Comp Meths Appl Mech Eng* 195(4–6):406–429
22. Parés N, Díez P, Huerta A (2008) Bounds of functional outputs for parabolic problems. Part I: Exact bounds of the discontinuous galerkin time discretization. *Comp Meths Appl Mech Eng* 197(19–20):1641–1660
23. Parés N, Díez P, Huerta A (2008) Bounds of functional outputs for parabolic problems. Part II: Bounds of the exact solution. *Comp Meths Appl Mech Eng* 197(19–20):1661–1679
24. Rannacher R, Suttmeier FT (1997) A feed-back approach to error control in finite element methods: application to linear elasticity. *Comp Mech* 19(5):434–446
25. Prudhomme S, Oden JT (1999) On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Comp Meths Appl Mech Eng* 176(1–4):313–331
26. Sarrate J, Peraire J, Patera AT (1999) Finite element error bounds for nonlinear outputs of the Helmholtz equation. *Int J Numer Meth Fluids* 11(1):17–36
27. Oden JT, Prudhomme S (2001) Goal-oriented error estimation and adaptivity for the finite element method. *Comp Maths Appl* 41:735–756
28. Díez P, Calderón G (2007) Goal-oriented error estimation for transient parabolic problems. *Comp Mech* 39(5):631–646
29. Rüter M, Stein E (2007) On the duality of finite element discretization error control in computational newtonian and eshelbian mechanics. *Comp Mech* 39(5):609–630
30. Cast3m. <http://www-cast3m.cea.fr>. Last accessed: 8 August (2008)
31. Matlab. <http://www.mathworks.com/>. Last accessed: 8 August (2008)
32. Peraire J, Patera AT (1997) Bounds for linear-functional outputs of coercive partial differential equations: local indicators and adaptive refinement. In: Ladevèze P, Oden JT (eds) Workshop on new advances in adaptive computational methods in mechanics. Elsevier Science Ltd, Oxford, pp 199–216
33. Paraschivoiu M, Peraire J, Patera AT (1997) A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations. *Comp Meths Appl Mech Eng* 150(1–4):23–50
34. Ladevèze P (2007) Strict upper error bounds on computed outputs of interest in computational structural mechanics. *Comp Mech* 42(2):271–286
35. Chamoïn L, Ladevèze P (2007) Bounds on history-dependent or independent local quantities in viscoelasticity problems solved by approximate methods. *Int J Numer Meth Eng* 71:1387–1411